AD-A160 008

(12)

"END-OF-FISCAL YEAR" REPORT

Speech Recognition: Acoustic, Phonetic and Lexical

Office of Naval Research
Contract N00014-82-K-0727

covering the period
1 October 1984 - 30 September 1985

Submitted by:

Victor W. Zue

This document has been approved for public release and sale; its distribution is unlimited.

October 1, 1985

DTIC
ELECTE
OCT 7 1985
S
A

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Research Laboratory of Electronics
Cambridge, Massachusetts 02139

DTIC FILE COPY

85 10 04 077

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE AD-A160 008

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S) ARPA Order No. 4585
6a. NAME OF PERFORMING ORGANIZATION Research Laboratory of Electronics Massachusetts Institute of Technology 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Advanced Research Projects Agency
6c. ADDRESS (City, State and ZIP Code) 77 Massachusetts Avenue Cambridge, MA 02139	7b. ADDRESS (City, State and ZIP Code) 1400 Wilson Boulevard Arlington, Virginia 22217
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Office of Naval Research Mathematical and Physical Sciences Res. Progr 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-82-K-0727
8c. ADDRESS (City, State and ZIP Code) 800 North Quincy Street Arlington, Virginia 22217	10. SOURCE OF FUNDING NOS. PROGRAM ELEMENT NO. / PROJECT NO. NR 049-542 / TASK NO. / WORK UNIT NO.
11. TITLE (Include Security Classification) Speech Recognition: Acoustic, Phonetic and Lexical Knowledge	
12. PERSONAL AUTHOR(S) by Victor W. Zue	
13a. TYPE OF REPORT End-of-Fiscal Year Rep 13b. TIME COVERED FROM 10/1/84 TO 9/30/85	14. DATE OF REPORT (Yr., Mo., Day) October 1, 1985 15. PAGE COUNT 5
16. SUPPLEMENTARY NOTATION	
17. COSATI CODES FIELD / GROUP / SUB. GR.	18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Work by Victor W. Zue and his collaborators is summarized here.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED X SAME AS RPT ☐ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Kyra M. Hall RLE Contract Reports	22b. TELEPHONE NUMBER (Include Area Code) (617)253-2569 22c. OFFICE SYMBOL

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE.

SECURITY CLASSIFICATION OF THIS PAGE

"End-of-Fiscal Year" Letter

1. Contract Information

(a) **Title:** Speech Recognition: Acoustic, Phonetic, and Lexical Knowledge Representation

(b) **ONR Contract No.:** N00014-82-K-0727

(c) **Principal Investigator:** Victor W. Zue

(d) **ONR Scientific Officer:** Cmdr. Allen Sears, DARPA-IPTO

(e) **Period Covered:** 1 October 1984 to 30 September 1985

2. Research Description

(a) Description of Research:

Our long-term research goal is the development and implementation of speaker-independent continuous speech recognition systems. It is our conviction that proper utilization of speech-specific knowledge is essential for advanced speech recognition systems. With this in mind, we have continued to make progress on the acquisition of acoustic-phonetic and lexical knowledge.

(b) Significant Results in the Last Year:

- We have completed the development of a continuous digit recognition system. The system was constructed to investigate the utilization of acoustic phonetic knowledge in a speech recognition system. Some of the significant development of this study includes a "soft-failure" procedure for lexical access, and the discovery of a set of acoustic-phonetic features for verification.
- We have completed a study of the constraints provided by lexical stress on word recognition. We found that lexical stress information alone can, on the average, reduce the number of word candidates from a large dictionary by more than 80%. In conjunction with this study, we successfully developed a system that automatically determines the stress pattern of a word from the acoustic signal.
- We have performed an acoustic study on the characteristics of nasal consonants and nasalized vowels. We have also developed recognition algorithms for nasal murmurs and nasalized vowels in continuous speech.
- We have finished the preliminary development of a system that aligns a speech waveform with the corresponding phonetic transcription.

(c) Plans for Next Year's Research:

- We will continue our investigation of the constraints imposed by the language on permissible sound patterns. In particular, we will investigate in greater detail the role played by lexical stress. We will also expand our study of lexical constraints to include larger linguistic units such as syllables and metrical feet.
- We will incorporate the knowledge gained through these studies into recognition systems. Specifically, we will try to improve the performance of the lexical stress recognition system, and we will continue our development of a broad phonetic classifier, with emphasis on handling continuous speech.

- We will investigate the feasibility of constructing a knowledge-based system that attempts to mimic the procedure of spectrogram reading.

(d) Presentations:

A. M. Aull and V. W. Zue, "Lexical Stress and Its Application in Large Vocabulary Speech Recognition," 108th Meeting of the Acoustical Society of America, Minneapolis, MN, Oct. 1984.

J. R. Glass and V. W. Zue, "An Acoustic Study of Nasal Consonants in American English," 108th Meeting of the Acoustical Society of America, Minneapolis, MN, Oct. 1984.

A. M. Aull and V. W. Zue, "Lexical Stress and Its Application in Large Vocabulary Speech Recognition," ICASSP 85, Tampa, FL, March 1985.

J. R. Glass and V. W. Zue, "Detection of Nasalized Vowels in American English," ICASSP 85, Tampa, FL, March 1985.

K. Moore and V. W. Zue, "The Effect of Speech Rate on the Application of Low-Level Phonological Rules in American English," 109th Meeting of the Acoustical Society of America, Dallas, TX, April 1985.

V. W. Zue and D. S. Cyphers, "The MIT *Spire* System," Speech Tech '85, New York, NY, April 1985.

V. W. Zue, "Human-Machine Communication by Voice: Computer Recognition of Speech," Annual Meeting of American Association for the Advancement of Science, Los Angeles, CA, May 1985.

J. R. Glass and V. W. Zue, "Analysis and Recognition of Nasal Consonants in American English," IASTED Robotics and Automation Conference, Lugano, Switzerland, June 1985.

H. C. Leung and V. W. Zue, "Automatic Alignment of Phonetic Transcriptions with Continuous Speech," IASTED Robotics and Automation Conference, Lugano, Switzerland, June 1985.

(e) Publications:

Papers

A. M. Aull and V. W. Zue, "Lexical Stress and Its Application in Large-Vocabulary Speech Recognition," *J. Acoust. Soc. Amer.*, suppl. 1, vol. 76, Fall 1984, p. S47.

A. M. Aull and V. W. Zue, "Lexical Stress and Its Application in Large Vocabulary Speech Recognition," *Proc. ICASSP 85*, vol. 4, pp. 1549–1552, 1985.

J. R. Glass and V. W. Zue, "Acoustic Characteristics of Nasal Consonants in American English," *J. Acoust. Soc. Amer.*, suppl. 1, vol. 76, Fall 1984, p. S1[illegible].

J. R. Glass and V. W. Zue, "Detection of Nasalized Vowels in American English," *Proc. ICASSP 85*, vol. 4, pp. 1569–1572, 1985.

V. W. Zue and D. S. Cyphers, "The MIT *Spire* System," *Proc. Speech Tech '85*, pp. 277–279, 1985.

Theses

F. R. Chen, "Acoustic-Phonetic Constraints in Continuous Speech Recognition: A Case Study Using the Digit Vocabulary," Ph.D. thesis, June 1985.

D. S. Cyphers, "Spire: A Research Tool," S.M. thesis, May 1985.

J. R. Glass, "Nasal Consonants and Nasalized Vowels: An Acoustic Study and Recognition Experiment," S.M. thesis, December 1984.

H. C. Leung, "A Procedure for Automatic Alignment of Phonetic Transcriptions with Continuous Speech," S.M. thesis, January 1985.

(f) Participants:

Mark Anderson
Francine R. Chen (Ph.D., June 1985)
D. Scott Cyphers (S.M., May 1985)
James R. Glass (S.M., December 1984)
Daniel Huttenlocher
Lori F. Lamel
Niels Lauritzen (S.M., May 1985)
Hong C. Leung (S.M., January 1985)
Kimberly Moore
Jacqueline Vaissiere
Victor W. Zue

Accession For
NTIS CRA&I ☒
DTIC TAB ☐
Unannounced ☐
Justification
By
Distribution/
Availability Codes
Dist | Avail and/or Special
A-1

(g) Other Sponsored Research:

- Title: Acoustic-Phonetics Based Speech Recognition
  Sponsor: Naval Electronic Systems Command (DARPA)
  Amount: $463,284.00
- Title: Speech Database Development
  Sponsor: Naval Electronic Systems Command (DARPA)
  Amount: $501,194.00
- Title: Tools for Speech Analysis and Research
  Sponsor: Space and Naval Warfare Systems Command (DARPA)
  Amount: $641,526.00

DTIC COPY INSPECTED

DISTRIBUTION LIST

	DODAAD Code	
Head Information Sciences Division Office of Naval Research 800 North Quincy Street Arlington, Virginia 22217	N00014	(1)
Administrative Contracting Officer E19-628 Massachusetts Institute of Technology Cambridge, Massachusetts 02139		(1)
Director Naval Research Laboratory Washington, D.C. 20375 Attn: Code 2627	N00173	(1)
Defense Technical Information Center Bldg. 5, Cameron Station Alexandria, Virginia 22314	S47031	(12)